TRANSFORMATION OF THE FUNDAMENTAL RELATIONSHIPS IN SEQUENTIAL ANALYSIS[1]

By H. Blasbalg[2]

Electronic Communications, Inc., Baltimore, Maryland

0. Summary. For the class of distribution functions given by

$$dP(x, \theta) = \exp [r(\theta)A(x) + s(\theta)B(x)]\, dw(x),$$

it is shown that a set of three transformations can be introduced which completely define the Sequential Probability Ratio Test for testing a hypothesis H_0 against H_1. When the observer specifies the threshold parameters θ_0 and θ_1 corresponding to the hypotheses H_0 and H_1 and the strength α, β of the test, he specifies the three transformations and hence the Sequential Test. However, there is an infinity of sets of parameter points $(\theta_0, \theta_1, \alpha, \beta)$ which satisfy the same transformations and hence define the same Sequential Test. The Operating Characteristic Function and the Average Sample Number Function are derived in terms of these transformations.

1. Introduction. Every pair of distributions leads to a two-parameter family of Sequential Probability Ratio Tests. A one-parameter family of distributions leads to a two-parameter family of probability ratios, and hence, one might expect a four-parameter family of Sequential Probability Ratio Tests. This is typically the case. In this paper it will be shown that there is a class of one-parameter families of distributions, each of which generates only a three-parameter family of Sequential Probability Ratio Tests. This includes the well-known exponential class, of which the best known examples are the normal family of unknown variance and known mean, the Bernoulli Distribution, and the Poisson Distribution.

Originally, the author proved that the well-known one-parameter family of exponential distributions

$$P(x, \theta) = v(\theta)w(x) \exp \{x\ell(\theta)\} \tag{1.1}$$

gives rise to this property. This has been recognized implicitly by Girshick [4]. Heuristic arguments supplied by L. J. Savage [3] lead to a more general one-parameter family of exponential distributions which also exhibit this property. Briefly, Savage's arguments indicate:

(1) if $r(x)$ and $s(x)$ are logarithms of probability ratios such that

$$\sum_{i=1}^{n} r(x_i) \leqq A \leftrightarrow \sum_{i=1}^{n} s(x_i) \leqq B \qquad \text{for all } n,$$

Received June 4, 1956; revised June 19, 1957.

[1] This research was supported by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command.

[2] Formerly with the Johns Hopkins University Radiation Laboratory, Baltimore, Maryland.

then,

$$r(x) = \lambda s(x),$$
$$B = \lambda A,$$

where λ is a constant.

Bringing the results of (1) to bear on the problem of interest, he concludes that:

(2) if a one-parameter family of distributions $P(x, \theta)$ gives rise to a three-parameter family of Sequential Probability Ratio Tests then

$$dP(x, \theta) = \exp [r(\theta)A(x) + s(\theta)B(x)]\, dw(x). \tag{1.2}$$

This family of distributions obviously includes the family in Eq. (1.1). The author's original proof can just as easily be applied to this more general family of distributions. We will therefore prove that the one-parameter family of exponential distributions given in Eq. (1.2) gives rise to only a three-parameter family of Sequential Probability Ratio Tests.

A Sequential Probability Ratio Test for testing a hypothesis H_0 against the alternative hypothesis H_1 (where H_0 and H_1 are mutually exclusive) for a given a priori probability ζ is defined by a set of four parameters α, β, θ_0 and θ_1. The hypothesis H_0 is accepted if $\theta \leqq \theta_0$ and H_1 is accepted if $\theta \geqq \theta_1$, $(\theta_1 > \theta_0)$. Also α is the probability of accepting H_1 when H_0 is true (type I error) and β is the probability of accepting H_0 when H_1 is true (type II error). These parameters define the Sequential Test and hence its fundamental characteristics, the Average Sample Number Function (ASN Function) and the Operating Characteristic Function (OC Function).

2. To obtain the statistical decision regions. Consider the one-parameter family of statistical distributions given by

$$dP(x, \theta) = \exp [r(\theta)A(x) + s(\theta)B(x)]\, dw(x). \tag{2.1}$$

To construct an example of such a family $A(x)$, $B(x)$ can be almost arbitrary functions and $w(x)$ an arbitrary measure on any measure space with the condition that $w(x)$ must be non-negative. A little care must be taken to insure that there will be a reasonably large class of pairs of numbers (r, s) such that,

$$\int \exp [r(\theta)A(x) + s(\theta)B(x)]\, dw(x) = 1. \tag{2.2}$$

The pairs (r, s) which normalize (2.2) form a smooth curve and $r(\theta)$, $s(\theta)$ can be parametric equations which define the curve with the understanding that different values of the parameter θ correspond to different points on the curve. Among the important distributions contained in this class are the Bernoulli, Poisson, and Gaussian.

In Sequential Analysis the observable upon which a decision is made is given by the logarithm of the probability ratio (likelihood ratio) which in this case is,

$$z(n) = [r(\theta_1) - r(\theta_0)] \sum_{i=1}^{n} A(x_i) + [s(\theta_1) - s(\theta_0)] \sum_{i=1}^{n} B(x_i). \tag{2.3}$$

The decision regions in Sequential Analysis are defined by the following: accept the hypothesis H_1 when

$$z(n) \geqq \log \frac{1-\beta}{\alpha} + \log \frac{\zeta}{1-\zeta} \tag{2.4}$$

and accept the hypothesis H_0 when

$$z(n) \leqq \log \frac{\beta}{1-\alpha} + \log \frac{\zeta}{1-\zeta}, \tag{2.5}$$

where

α = probability of accepting H_1 when H_0 is true ($\theta \leqq \theta_1$),
β = probability of accepting H_0 when H_1 is true ($\theta \geqq \theta_1$),
ζ = a priori probability of the hypothesis H_0.

We now substitute Eq. (2.3) into Eqs. (2.4) and (2.5) and obtain

$$\sum_{i=1}^{n} A(x_i) + \frac{s(\theta_1) - s(\theta_0)}{r(\theta_1) - r(\theta_0)} \sum_{i=1}^{n} B(x_i) \geqq \frac{\log \frac{1-\beta}{\alpha} + \log \frac{\zeta}{1-\zeta}}{r(\theta_1) - r(\theta_0)} \tag{2.6}$$

corresponding to H_1 and

$$\sum_{i=1}^{n} A(x_i) + \frac{s(\theta_1) - s(\theta_0)}{r(\theta_1) - r(\theta_0)} \sum_{i=1}^{n} B(x_i) \leqq \frac{\log \frac{\beta}{1-\alpha} + \log \frac{\zeta}{1-\zeta}}{r(\theta_1) - r(\theta_0)} \tag{2.7}$$

corresponding to H_0. Let us now define the following three transformations:

$$a = \frac{\log \frac{\zeta}{1-\zeta} + \log \frac{1-\beta}{\alpha}}{r(\theta_1) - r(\theta_0)}, \tag{2.8}$$

$$b = \frac{\log \frac{\zeta}{1-\zeta} + \log \frac{\beta}{1-\alpha}}{r(\theta_1) - r)\theta_0)}, \tag{2.9}$$

$$c = \frac{s(\theta_1) - s(\theta_0)}{r(\theta_1) - r(\theta_0)}. \tag{2.10}$$

Substituting (2.8), (2.9), and (2.10) into (2.6) and (2.7) yields

$$\sum_{i=1}^{n} A(x_i) + c \sum_{i=1}^{n} B(x_i) \geqq a \tag{2.11}$$

for the acceptance of H_1 and

$$\sum_{i=1}^{n} A(x_i) + c \sum_{i=1}^{n} B(x_i) \leqq b. \tag{2.12}$$

It is therefore seen that the a, b, c transformations completely define the decision regions of the Sequential Test for the one-parameter family of distributions considered. Furthermore, for a given a priori probability ζ, there is an infinity of sets of values $(\theta_0, \theta_1, \alpha, \beta)$ which yield the same decision regions.

For a given a priori probability ζ, it is known that the Sequential Probability Ratio Test is optimum [1] in the sense that the ASN Function at the parameter point $(\theta_0, \theta_1, \alpha, \beta)$ is less or equal to the ASN Function for any other Sequential Test. Since these parameters specify the a, b, c transformations of the Sequential Probability Ratio Test, this test is also optimum for a given set a, b, c. For a given a priori probability ζ and a given set a, b, c there exists an infinity of parameter points $(\theta_0, \theta_1, \alpha, \beta)$ which satisfy the transformations. We therefore conclude that a given Sequential Test is optimum at an infinity of parameter points $(\theta_0, \theta_1, \alpha, \beta)$ which satisfy the given transformations.

One can easily express Wald's approximations to the OC and ASN Functions in terms of the a, b, c transformations. The OC Function can be obtained by solving Eq. (2.8), (2.9), and (2.10) for the appropriate variables, substituting these into the parametric equations which define the OC Function [2] and then introducing a new transformation,

$$u = \exp (h[r(\theta_1) - r(\theta_0)]). \tag{2.13}$$

As h ranges over the entire real line, u ranges over the positive half of the real line. The parametric equations for the OC Function are then given by

$$L(u) = \frac{u^a - 1}{u^a - u^b} \tag{2.14}$$

and

$$E_\theta[u^{A(x)+cB(x)}] = 1. \tag{2.15}$$

At the indeterminant point $u = 1$,

$$L(1) = \frac{a}{a + |b|}. \tag{2.16}$$

The point $u = 1$ corresponds to the value $\theta = \theta'$ for which

$$E_\theta[A(x)] + cE_\theta[B(x)] = 0.$$

In a similar manner the ASN Function [2] can be shown in terms of the new transformations as

$$E_\theta(n) = \frac{bL(u) + a[1 - L(u)]}{E_\theta\{A(x)\} + cE_\theta\{B(x)\}}. \tag{2.18}$$

When $u = 1$, $\theta = \theta'$ and

$$E_{\theta'}(n) = \frac{ab}{E_{\theta'}[A(x) + cB(x)]^2}. \tag{2.18}$$

The author wishes to express his thanks to L. J. Savage for his interest in this problem and his many useful suggestions.

REFERENCES

[1] D. Blackwell and M. A. Girshick, *Theory of Games and Statistical Decisions*, John Wiley and Sons, Inc., New York, 1954.
[2] A. Wald, *Sequential Analysis*, John Wiley and Sons, Inc., New York, 1947.
[3] L. J. Savage, "When different pairs of hypotheses have the same family of likelihood-ratio," this issue of these *Annals*.
[4] M. A. Girshick, Contributions to the Theory of Sequential Analysis, II, III, *Ann. Math. Stat.* Vol. 17 (1946), pp. 282–298.

WHEN DIFFERENT PAIRS OF HYPOTHESES HAVE THE SAME FAMILY OF LIKELIHOOD-RATIO TEST REGIONS[1]

By Leonard J. Savage

University of Chicago

Blasbalg [1], in this issue of these *Annals*, shows that certain families of distributions are especially simple, or degenerate, from the point of view of sequential tests. The main object of this note is to show briefly that these are (at least practically) the only families thus degenerate; some preliminary and related conclusions are also demonstrated.

Let F and G be a pair of probability measures on a space X with elements x, and let ℓ be the logarithm of the likelihood ratio of F with respect to G. ℓ is of course defined only mod $(F + G)$, that is, only up to sets simultaneously of F and G measure 0. If x_i is a sequence of values of x, then a likelihood-ratio critical region in X^n is defined by

$$R(A, n) = \left\{(x_1, \cdots, x_n) : \sum_1^n \ell(x_i) \leqq A\right\}. \tag{1}$$

The innocuous ambiguity of ℓ of course induces corresponding ambiguity in R.

This family of sets R is simplest to study when the distribution of ℓ is non-

Received May 3, 1957.

[1] Research was carried out at the Statistical Research Center, University of Chicago, under sponsorship of the Statistics Branch, Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

I wish to thank H. Blasbalg for his help in coordinating his paper [1] and this note, and I thank W. H. Kruskal for many suggestions. Before writing this, I had the privilege of seeing a related manuscript based on an idea of M. A. Girshick.